A COMPARATIVE STUDY OF DIFFERENT TYPES OF MOTHER WAVELETS FOR HEARTBEAT BIOMETRIC VERIFICATION SYSTEM

By

LEE JIAN HAN

A Dissertation submitted for partial fulfilment of the requirement for the degree of Master of Science (Electronic Systems Design Engineering)

August 2017

Acknowledgements

First and foremost, I would like to sincerely thank my supervisor, Associate Professor Dr. Dzati Athiar binti Ramli for the guidance and direction provided in this research project. Without her generous help, this research would not be possible. Her attention, enthusiasm and patience that were presented while guiding me in the research is much appreciated.

Next, I would like to thank Keysight Technologies for providing me the opportunity and support in pursuing this research project. I am also grateful to the PPKEE department at Universiti Sains Malaysia for their teachings, technical advice and inspiration to begin this research project. This work was supported by Universiti Sains Malaysia under the Fundamental Research Grant Scheme (6071266).

Finally, I would like to thank my family and friends for their continuous support during my studies.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abbreviations	Expansion
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
EER	Equal Error Rate
EMD	Empirical Mode Decomposition
FAR	False Acceptance Rate
FRR	False Rejection Rate
GAR	Genuine Acceptance Rate
HRV	Heart Rate Variability
KPCA	Kernel Principal Component Analysis
LDA	Linear Discriminant Analysis
MLP	Multi-Layer Perceptron
MLSP	Multiscale Local Shape Patterns
OAA	One Against All
PAR	Pulse Active Ratio
PAW	Pulse Active Width
PWM	Pulse Width Modulation
ROC	Receiver Operating Characteristic
SVM	Support Vector Machine

KAJIAN PERBANDINGAN TERHADAP PELBAGAI JENIS IBU WAVELET UNTUK SISTEM PENGESAHAN BIOMETRIK DENYUTAN JANTUNG

Abstrak

Kebelakangan ini, teknologi biometrik terkini sedang beralih arah kepada penggunaan isyarat elektrokardiogram (ECG) sebagai satu modaliti terbaru untuk sistem pengesahan. Isyarat ECG mengandungi maklumat yang mencukupi untuk mengesahkan individu kerana ia adalah unik kepada setiap individu. Satu daripada teknik yang dipercayai yang mampu mengektrak maklumat penting daripada isyarat ECG ialah dengan menggunakan transformasi wavelet. Namun, terdapat satu cabaran dalam mengimplimentasikannya kerana perbezaan jenis dan peringkat ibu wavelet akan mempengaruhi prestasi pengesahan. Jadi, dalam kajian ini, satu kajian perbandingan akan dibuat bagi menyemak jenis dan peringkat ibu wavelet yang optimum yang boleh mewakili ciri yang terbaik untuk sistem verifikasi ini. Tiga jenis ibu wavelets iaitu Symlet, Daubechies dan Coiflet dengan turutan peringkat dari satu kelima telah dikaji di dalam kajian ini. Ciri yang diektrak kemudiannya dilatih dengan pengelas Mesin Vektor Sokongan (SVM) bagi membina model untuk mengesahkan ciri yang diektrak. Prestasi sistem pengesahan biometrik ECG ini dinilai dengan Penerima Operasi Ciri (ROC) dan Kadar Ralat Sama (EER). Keputusan kajian menunjukkan sistem yang telah dibangunkan ini menghasilkan prestasi yang terbaik apabila ibu wavelet Coiflet peringkat ketiga digunakan sebagai ciri dengan prestasi EER bernilai 10.755% diperolehi.

A COMPARATIVE STUDY OF DIFFERENT TYPES OF MOTHER WAVELETS FOR HEARTBEAT BIOMETRIC VERIFICATION SYSTEM

Abstract

Recently, advanced biometric technology is turning to the use of electrocardiograms (ECG) signal as new modality for verification system. The ECG signal contains sufficient information to verify an individual as it is unique to everyone. One of the feasible methods to extract the salient information from ECG signal for verification is by using wavelet transform. However, there is a challenge in implementing it as different types and orders of mother wavelet used will yield different verification performance. Therefore, in this study, a comparative study is done so as to investigate the optimum type and order of mother wavelet that represents the best feature for the verification system. Three different types of mother wavelets i.e. Symlet, Daubechies and Coiflet with order ranging from one to five have been studied in this research. The extracted features are then trained by using SVM classifier to generate a model to verify the features. The performance of the ECG biometric verification system is evaluated with the Receiver Operating Characteristic (ROC) plot and Equal Error Rate (EER). Experimental result showed that the developed system achieves the best performance when the 3rd order Coiflet is used as feature with an EER score of 10.755% is achieved.

CHAPTER 1

INTRODUCTION

1.1 Background

Living in an era of advanced security technology where there is a need of verification system that can help to protect the assets of millions of people from being stolen is of fundamental importance. The most basic security system consists of verification system which would require some form of identification where a security password is required to access the data that were kept. The automatic verification system is very popular due to its reliability that can ease people in gaining access to their data securely.

There are many types of verification systems being used today to verify an authorized person. These systems authenticate such a person with the reliance of stored information in the form of ID card, username, token and many more. Besides that, majority of the businesses today depend on electronic data and internet connection for their operations which lead to many personal and business information being transferred and stored online. With such an advanced technology, this sensitive information is rather easily accessed and obtained by scammers. These scammers use various methods which are specifically crafted to obtain the verification information desired.

According to a research on identity fraud conducted in 2016, a total of 13.1 million victims of identity fraud had loss a staggering amount of USD 15 billion in the United State alone [1]. This astounding figure showed that there is a need of improvement in security technology to enhance the automatic verification system. One of the verification methods used by verification system is human biometrics. Recently, there has been a lot of biometric studies regarding cardiovascular signals being used in verification system. These signals are generated from cardiovascular activity which are measured in the form of electrical signals; are called as electrocardiograms (ECG). ECG signals varies by person due to the unique anatomical structure of the heart.

Furthermore, ECG signals are able to tell the differences between a normal and an abnormal behavior of the cardiovascular activities. Normal behavior ECG signals will show that a person is in healthy condition while the abnormal behavior shows that a person is unhealthy. Besides that, these ECG signals also contain sufficient information for use in verification and is robust enough to prevent circumventions and threats from attackers. Thus, ECG based biometric verification system is a reliable security system. In the future, when the sensors for ECG are becoming portable and stable, it shall be implemented in verification system for businesses usage.

1.2 Problem Statements

The new generation of biometric verification system that uses ECG signal which is known for its medical usage has been very popular recently. The use of ECG signal for biometric verification is due to the liveness detection, robustness to attack, universality and permanence [2]. There are many ECG biometrics verification systems that has been researched in the past decades which are proven successfully in human verification. These biometrics verification system are developed with various method used in feature extraction and verification.

Despite the various method used to developed the ECG biometric verification system, there are still many challenges faced when developing the ECG biometrics verification system using wavelet transform for feature extraction and Support Vector Machine (SVM) for verification. This verification system provides an ambiguity on the performance when wavelet transform is used for feature extraction together with SVM classifier. The wavelet transforms of the ECG signal will give different performance based on the types and orders of the mother wavelet used. There is no certain that which types and orders of mother wavelet used in wavelet transform that determines the best performance for an ECG biometric verification system.

Hence, the ECG biometric verification system needs to be developed with wavelet transform as feature extraction together with the verification from the SVM for further analysis. Besides that, the use of SVM classifier's score will help determine the performance of the ECG biometric verification system. Then only the comparative study on the different types and orders of mother wavelet shall be performed. From the comparative study on the different types and orders of mother wavelet used, the mother wavelets that provide the best performance of the verification system shall be determined.

1.3 Research Objectives

The main aim of this research is to compare the performance of different types of mother wavelets for ECG biometric verification. In order to do this, the following objectives must be achieved.

1) To developed feature extraction algorithms based on Symlet, Daubechies and Coiflet as mother wavelets.
2) To classify the ECG features using Support Vector Machine (SVM) classifier.
3) To evaluate the performance of the system based on Equal Error Rate (EER), False Acceptance Rate (FAR), False Rejection Rate (FRR) and Genuine Acceptance Rate (GAR).

1.4 Research Scope

This research will focus on ECG based biometric verification system that uses wavelet transform for feature extraction and SVM for classification of data. The use of the ECG signal to be evaluated in this research is limited to only the normal signal. Hence, the ECG screening algorithm will be developed in this study in order to exclude the abnormal ECG signal from the database. The normal signal is collected from a healthy individual and with the individual at resting position. While the abnormal signal is collected from unhealthy individual and with individual not at resting position. The source of the ECG data is taken from the ECG-ID database (ECGIDDB) which is hosted at PhysioNet website. Moreover, the research will also focus on the evaluation and comparative analysis of the performance of the biometric verification system which uses different mother wavelet and different orders of the mother wavelet.

1.5 Thesis Outline

The thesis is organized by dividing into 5 chapters as follows. The first chapter will cover the introduction for this thesis where it consists of problem statement, research objectives and research scope. The second chapter will present the literature review where research and studies related to ECG biometric verification system done by others are documented. This chapter also includes basic ECG signal explanations, wavelet transform theory as feature extraction and SVM as the classifier. The third chapter will describe the methodology in detail and the method implemented in this project. The fourth chapter will present the analysis and discussion of the preprocessing, feature extraction and classifier results. Moreover, Chapter 4 describes the performance of numerous mother wavelet and the orders of mother wavelets used. Lastly, Chapter 5 will conclude the thesis.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

In this chapter, numerous researches regarding ECG biometric verification system shall be discussed. This will lead to some of the challenges faced by the researcher on the techniques used for feature extraction and the models used for classification of the data collected. There are three types of feature extraction, one being fiducial based, second is the non-fiducial based and third is a hybrid where both fiducial and non-fiducial based are combined. These types of feature extraction methods were adapted into the studies and are crafted specifically for ECG signal. Besides that, the classifier used for ECG biometrics is widely different among the studies. Some of the studies use simple classifier while other uses more complex classifiers.

The following section shall discuss the fundamentals and theories used in ECG biometric verification system. The fundamentals of ECG signal are evaluated and explained in further details. One of the theories used in this research is the wavelet transformation where the algorithm of the transformation is explained. Furthermore, the classifiers used in this research which is a SVM classifier will be discussed as well.

2.2 ECG Signal

An ECG signal is the measurement of the heart movement in terms of electrical activity. ECG signal is as practical and secure as compared to other biometric data due to the ECG signal are easily measurable with today’s technology and are very hard to be replicated. ECG signals are complex and difficult to duplicate due to the signal being generated from the biological nature of the heart. The ECG signal carries adequate information about the movement of the heart and is very personalize as the structure of an individual’s heart is uniquely developed.

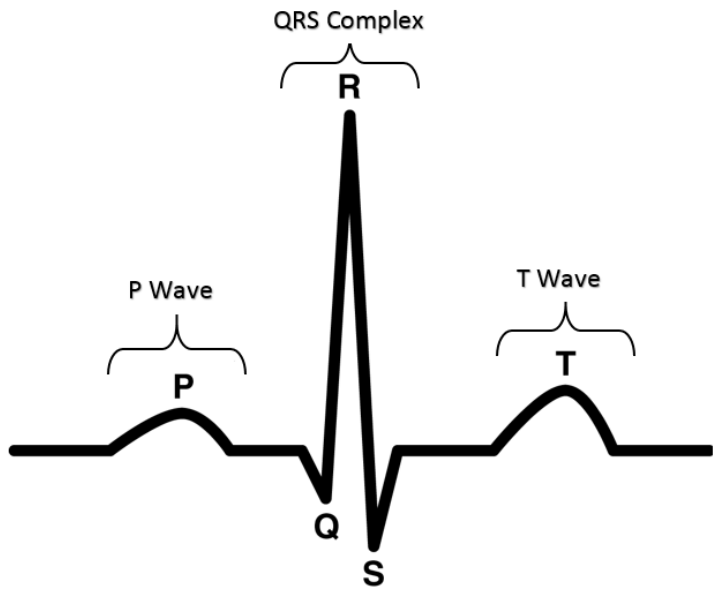

Figure 2.1: Segment of continuous ECG signal.

A regular ECG signal is plotted in the time domain versus the amplitudes. An ECG signal is separated into three main components which are the P wave, QRS complex and T wave as shown in Figure 2.1 above. An ECG signal always starts with the P wave where the atrial depolarizes follow by the QRS complex due to the depolarization of the ventricular area and then completes the cycle with the T wave which is due to the ventricular repolarization [3].

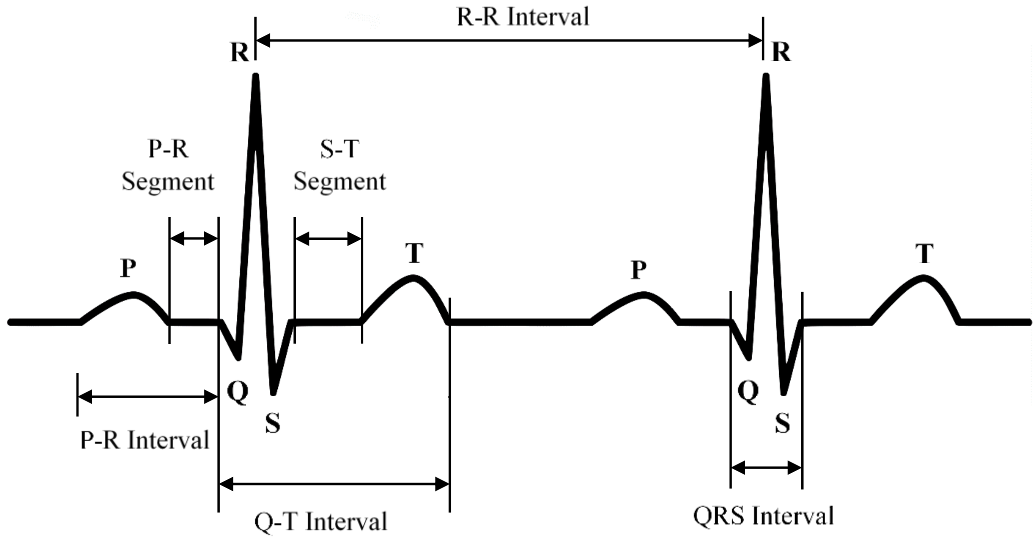

Figure 2.2: In-depth detail of an ECG signal.

An in-depth detail of the ECG signal is as shown in Figure 2.2 above. There are several parameters on an ECG signal that can be measured in terms of time. One of the parameters would be the P-R interval which is the period from the start of the P wave until the start of the QRS complex. The P-R interval has a typical duration of 120 milliseconds, ms. While for the P-R segment, the duration is at the end of P wave to the start of QRS complex and is typically flat. The next parameter would be the Q-T interval which is the period from the start of QRS complex to the end of the T wave and have a duration of less than 400 ms.

The S-T segment parameter is between the end of QRS complex until the end of T wave and it represents the ventricular depolarization. Another parameter would be the QRS interval which has a typical duration of 70 to 110 ms and the interval is between the end of the P-R interval to the start of S-T segment [2]. The last parameter would be the R-R interval where the interval is between the R peak of one ECG pulse and the preceding

ECG pulse. The R-R interval can also represent the variability of the heart period and are normally used for medical studies.

Identity verification from an ECG signal recording can be categorized by using various method such as the number of ECG data channel used, operational setting, the method to generate desired features and type of classifier used [4]. The number of ECG data channel used typically are one to three channels or 12 lead ECG signal. Most of the study in ECG biometrics uses single channel ECG signals as the single channel ECG signal provides sufficient information to be used for biometric verification [5]. While some of the studies use more than one channel ECG signals which also provides adequate information for biometric verification. The more channel used of an ECG signal does not mean it is necessarily better as more channel ECG signal requires more processing time where extra steps are required to be performed in order to combine scheme of the channel effectively.

Besides that, the ECG signal can also be categorized based on how the ECG signal is collected and what is the condition of an individual when the ECG signals are collected. There is a standard placement of electrode to collect the ECG signals for studies while there are also nonstandard placements of the electrode being used for studies as well. The nonstandard placement includes the placement of pads on the individual’s thumbs, palms [6], and some electrode placed on the wrist together with the finger [7]. Normally ECG signals are taken from the individual in resting conditions and are used for studies in the biometric verification.

2.3 Wavelet Transform Theory

A raw ECG signal does not provide any further information for biometrics purposes and the ECG signal needs to be transformed by any mathematical transformation to unlock further information contains in the ECG signal. Wavelet transform is being used as the mathematical transformation due to the capability to provide a time-frequency domain of the ECG signal. Wavelets transform can accurately deconstruct and reconstruct finite, non-periodic and non-stationary signals. Besides that, the wavelet transform can represent a function with sharp peaks and discontinuities [8].

The wavelet transform can be determined by plotting a function $f(t)$ in the space of square integral functions, $L^2(R)$ to another signal $W_f\,(a,\ b)$ in L^2R^2 where $(a,\ b)$ are continuous with a representing the scaling parameter and b representing the shifting parameter. The wavelet transform basis in time and frequency domain can be defined with the mother wavelet $\psi(t)$ and continuous scaling and shift parameters $(a,\ b)$. The wavelet transform will then give an equation as below [9].

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{2.1}$$

The continuous wavelet transforms (CWT) of continuous time $f(t)$ can be define as

$$W_f(a,b) = \langle f(t), \psi_{a,b}(t) \rangle \tag{2.2}$$

where wavelet equation (2.1) combined with equation (2.2) giving equation (2.3).

$$W_f(a,b) = \frac{1}{\sqrt{|a|}}\int_{-\infty}^{+\infty} f(t)\,\psi\left(\frac{t-b}{a}\right)\,dt \tag{2.3}$$

CWT have an advantage in the time-frequency resolution, however, the CWT is still unable to represent functions concisely. Thus, the discretized of the continuous scaling and shift parameters (a, b) to proper discrete values can define the discrete wavelet transform (DWT). Let $a = a_0^m, b = nb_0 a_0^m$ to get the basis of DWT as shown in equation (2.4).

$$\{\psi_{mn}(t)\} = a_o^{-m/2}\psi(a_0^{-m}t - nb_0) \quad m, n \in Z \tag{2.4}$$

Expressing any $f(t) \in L^2(R)$ in the superposition of equation (2.5) gives the DWT equation when the sets $\{\psi_{mn}(t)\}$ called frame is complete for some $\psi(t), a$ and b [9].

$$f(t) = \sum_m \sum_n d_{m,n}\psi_{mn}(t) \tag{2.5}$$

where the DWT coefficients, $d_{m,n}$ is defined as

$$d_{m,n} = \langle f(t), \psi_{mn}(t)\rangle = \frac{1}{a_o^{-m/2}} \int f(t)\psi(a_0^{-m}t - nb_0)dt \tag{2.6}$$

A one-dimensional DWT scheme can be illustrated in the Figure 2.3. With a signal *s* of length *N*, the DWT consist of $\log_2 N$ stages max. The signal will first be decomposed simultaneously through a low and high pass filter of length 2*N* giving two filtered signal *F* and *G* respectively where the length of signal F and G are $n + 2N - 1$, if $n = length(s)$. Next the filtered signal *F* and *G* is down sampled to get the approximation coefficients, cA_1 and detail coefficients, cD_1 where cA_1 and cD_1 have a length floor $\left(\frac{n-1}{2}\right) + N$ [10].

The following step would be splitting the approximation coefficients cA_1 into two parts using the same scheme illustrated in Figure 2.3. The signal, *s* will be replaced by approximations coefficient, cA_1 in the scheme. This will produce another approximation coefficients, cA_2 and detail coefficients, cD_2 and so on if the scheme is repeated. The one-dimensional wavelet decomposition of a signal, s at level j will result in the following structure: $\left[cA_j, cD_j, \ldots, cD_1\right]$ [10]. An example of the structure containing level 4 is shown in Figure 2.4.

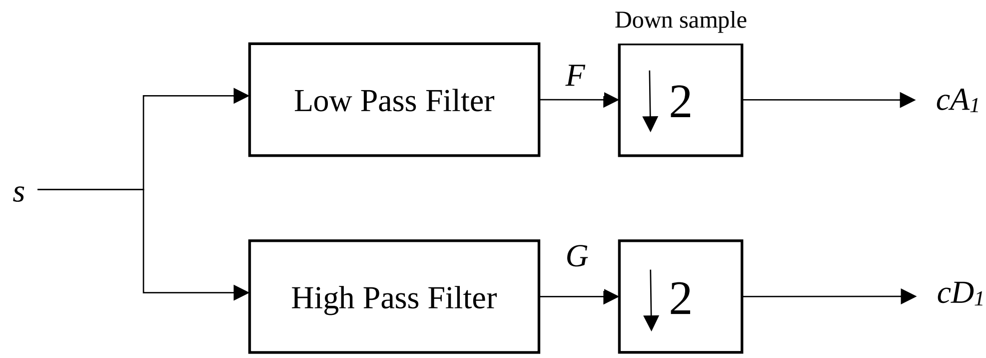

Figure 2.3: One-dimensional DWT scheme.

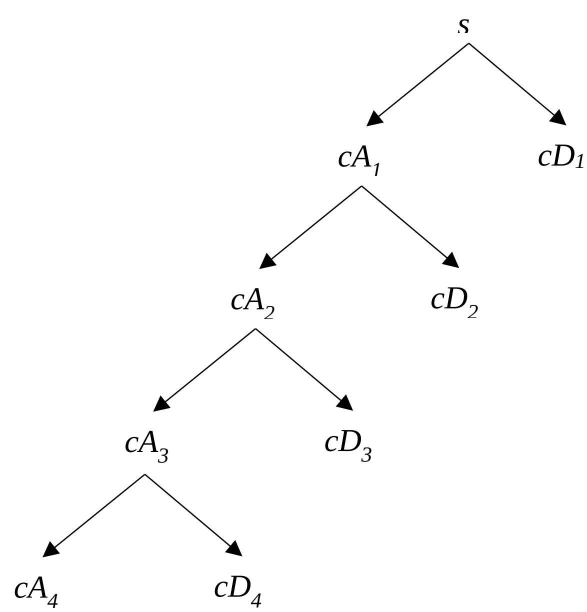

Figure 2.4: Structure of 4 level decomposition.

2.4 Support Vector Machine (SVM)

Normally, SVM are used for both classifications and regression challenges. SVM is a type of learning machine where it uses learning algorithms to learn from a set of training data and using those training data to build SVM models. The SVM models will then attempts to classify a set of new data into its respective classes. These SVM models performed the classification by generating a hyperplane that separates the data into two classes. Provided a training data, a set of input vectors can be generated where the term x_i represents each of the input vectors with several component features. Pairing the input vectors with its corresponding labels denotes as y_i and there will be *m* of such pair giving $(i = 1, \ldots, m)$.

In Figure 2.5, the depiction of the training data labelled as data points in an input space with a hyperplane can be seen. The hyperplane separating the data points to two sides where one side of the data points are labelled $y_i = positive\ (+)$ and the other side as $y_i = negative\ (-)$. The hyperplane generated by the SVM will have maximized margin in between the location of the separated labelled points. The support vectors are known as the closest points on both sides which have the most influence of the separating hyperplane. The linear separating hyperplane is as given in equation (2.7).

$$w.x + b = 0 \tag{2.7}$$

where b is the offset of the hyperplane from the origin in input space, *x* are points located within the hyperplane and are normal to the hyperplane and *w* is the weight which determines the orientation [11].

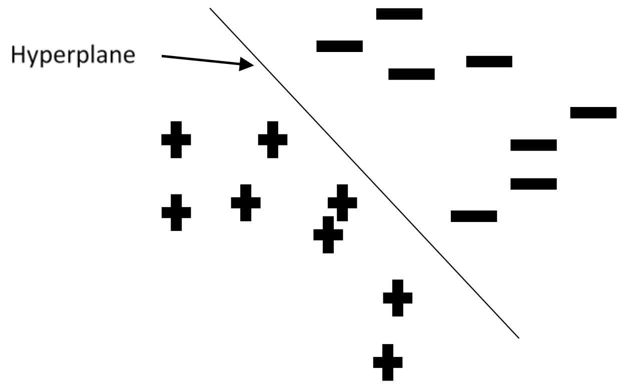

Figure 2.5: Linear classification of data.

In real-life, many data can come from various forms and may be complex to classify not like the linear classification shown in Figure 2.5 above. Data such as the extracted features from ECG signals are complex to classify and could not be accurately classified by just using linear classifier. With the introduction of a kernel function in SVM, it will help to transform the input space into a higher dimension space where the SVM can easily separate those complex labelled points.

The hyperplane generated by the SVM for input vector z, would be given as equation (2.8).

$$\phi(z) = \sum_{i=1}^{m} \alpha_i^* \, y_i K(x_i, z) + b^* \tag{2.8}$$

Where α_i^* is the Lagrange multipliers of the optimal value, the y_i is the labels of data points while the $K(x_i, z)$ is the kernel function and the b^* denotes the value of the bias at optimality [11].

2.5 Researches in ECG Biometrics

There are countless preceding works in the ECG biometric verification system which covers various methods to achieve good performance for its verification system. In this section, overviews of the type of feature extraction and classifier approaches that are used in the field of ECG biometric verification system found in the literature are presented.

Feature extractions from ECG signals can be classified into three types which are fiducial based, non-fiducial based and hybrid [4]. The fiducial based method of feature extractions relies on the extraction of amplitude, temporal, area, angle, or dynamically of a single pulse cycle of a heartbeat. Fiducial based requires characteristic points as a reference on the ECG signal to provide enough information for biometric verification. For the non-fiducial based method, characteristic points are not used as features from an ECG signal. Features like wavelet coefficients and autocorrelation coefficients are used instead while the hybrid method combines both fiducial and non-fiducial based features.

Furthermore, the types of classifiers used for ECG verification are based on the features where the classifiers used vary in different studies. Some of the studies use simple classifiers like nearest neighbor classifiers, nearest center classifiers, linear discriminant analysis (LDA) classifiers and others. While some of the studies use more complex classifier such as neural network classifier and support vector machines.

In the researches of ECG biometric verification system, one of research discussed by Hegde (2011) proposed a feature extraction technique that uses a non-fiducial method [12]. Radon transform is applied to the image of an ECG signal and Standardized Euclidean distance is applied on the Radon image to obtain the feature vector. There are

135 ECG signals taken from PhysioNet QT database and MIT-BIH database as the test subject. The classifier used here is correlation coefficient of more than 0.9 where the feature vector stored in the database is compared to the test ECG image and calculated the authenticity of a person. The performance of the methods achieves a FAR of 2.19% and FRR of 0.128%.

In this next study, a fiducial based feature extraction is proposed where the feature vector are based on the slopes and angles of the ECG signal [13]. The characteristics of the feature are from six slopes and three angles as shown in Figure 2.6. The database used contains 100 samples of ECG signals which are collected from volunteers. Such private database could not be used by others to study or perform further analysis and comparison of the method. The classifier used in this study is Multi-Layer Perceptron (MLP) which is a multilayer neural network and the performance achieved a verification rate of 96.44%. The performance evaluation is only the verification rate and the lack of other results like FAR and FRR which is common in biometric performance evaluation. This makes it hard to have a further comparison with this research.

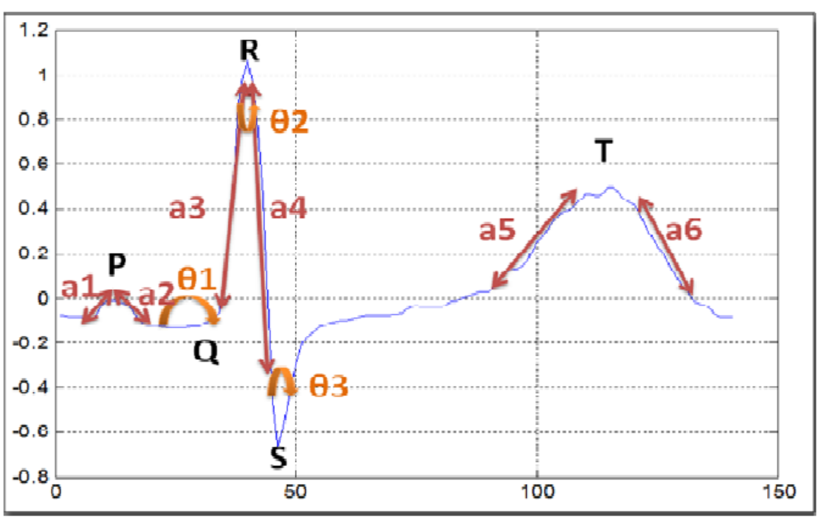

Figure 2.6: Six slopes and three angles that were selected [13].

Another research proposed a fiducial based feature extraction which is known by Pulse Active Width (PAW) where it is using the concept of Pulse Width Modulation (PWM) [14]. The ECG signal will be modulated with a triangular wave and resulted in the PAW which the modulated signal is controlled by the modulation index and modulation factor together with the maximum modulated amplitude set to 1. Euclidean distance is used in this study to verify the authenticity of the ECG signal where the value is generated from the differences between the feature vectors. The performance of the PAW is measured in Equal Error Rate (EER) at 0.1801 from 98 individual's data in PTB database.

In this following research, quite similar as the PAW research where the feature extraction uses the Pulse Active Ratio instead of comparing the modulated pulse width [15]. The ECG signal is still modulated with the triangular wave and this time the PAR feature vector is formed with a different set of parameters in the modulated signal which is the modulation index, modulation factor, maximum and minimum modulated signal amplitude. The resulted PAR or feature vectors are verified with Euclidean distance as well and achieve an EER of 0.1915 for Arrhythmia and 0.0989 for the healthy subject. The performance could not directly be compared with the PAW method due to a different set of databases being use. In this research 112 subjects are from PTB ECG database and out of the 112 subjects, 98 subjects have arrhythmia or irregular heartbeat and the rest are healthy subject.

The next research uses the non-fiducial method as the feature extractor which is the Autocorrelation method and Linear Discriminant Analysis (LDA) as the classifier for the biometric verification system [2]. There are 52 subjects used in this study and it achieves a verification rate of 81.48% for 5 seconds long ECG and 92.3% for 3 min recordings of ECG signal. However, the research also did use empirical mode decomposition (EMD) for recognizing the emotional patterns of the ECG signal and using it as biometric verification system. The reported results are 3.96% for EER with the implementation of biometric template updates.

In this research, the Heart Rate Variability (HRV) is used for generating 101 features and from there only 10 HRV features are reliable enough for biometric use [16]. The 101 features are generated using statistical, spectral, time frequency and nonlinear techniques which are a hybrid method of feature extraction. Furthermore, the 10 HRV features are selected based on five different algorithms which are the statistical dependency, mutual information, random subset feature selection, sequential forward selection and sequential floating forward selection. The R-R intervals of 81 subjects are collected and there is 11% of the subject having an unhealthy signal. The K-Nearest Neighbor classifier is used for classification and the performance achieves a verification rate of 82.22%.

The following research proposes an approach using a non-fiducial method where the ECG signal are segmented into five and multiscale local shape patterns (MLSP) are applied at each sample point [17]. The results of the MLSP give a pattern distribution which are used as feature vectors. In this study, ECG signals are taken from PTB database which there are 60 subjects chosen for the experiment. The Euclidean distance is used as a tool to verify the authenticity of the ECG signal which it achieves a verification rate of 91.67%.

The other research that uses non-fiducial method are using the discrete wavelet transform (DWT) to extract the feature vectors [6]. The mother wavelet used is Daubechies wavelet order 3 (DB3) with five level of decompositions. The feature vectors are from the DWT which is performed on a single ECG signal where there is an average of 100 ECG cycles. There is a total of 80 individual's ECG signals obtained from the PhysioNet database and all the ECG are from healthy individual. The use of Random Forest as classification resulted with a 100% verification rate. The random forest classifier depends on the correlation between the trees in the forest as the error rate will increase as the correlation between two trees increases. The training set needed to have low error rate for the random classifier to perform better.

In this research, the same non-fiducial based method are used as [6] which is the DWT but the mother wavelet used are Symlet, Daubechies and Coiflet [7]. There are no mentions of the mother wavelet order and level of decomposition used. In addition, the feature vectors are extracted from each of the segmented ECG signals which give more feature points as the training dataset for training the classifier. The ECG signals are collected from 50 individuals from the built heartbeat detection kit. The classifiers used is the SVM with optimization on the parameter and it achieves an EER of 2.0069% with GAR of 97% and FAR of 1%.

For this recent research, the DWT is used to denoise the ECG signal and the mother wavelet used is coif3 and decomposed at level 4 [18]. The non-fiducial based is used as feature extraction with combination of autocorrelation and Kernel Principal Component Analysis (KPCA). In this study, a total of 52 numbers of subjects are taken from volunteers and the signals are verified before using. The classifier used here is a Gaussian

(One Against All) OAA SVM where the SVM is using the Gaussian function as the kernel trick to achieve a non-linear classification. The results show that the ECG biometric verification achieves a performance with FRR of 4.83% and FAR of 3.5%.

In summary, the feature extractions which are fiducial based will depend on a specific incorporated location in the ECG signals for feature vectors and requires extra processing step to uncover the specific parameter. Furthermore, there are several anomalies which could affect the morphology of the signal significantly making the signal difficult to localize [6]. While for the non-fiducial based feature extraction, it will not depend on the morphology of the ECG signal and does not require any extra step in localizing the parameter from the ECG signal. This will also lead to improving the processing speed of uncovering the feature vectors.

Based on the literature reviewed, the non-fiducial based feature extraction that showed significant results is the DWT as in [6][7][18]. There are various types of mother wavelets and various order of mother wavelet being used in DWT in between those paper to obtain the features from the ECG signals. The biometric verification system is not solely dependent on the feature extraction but the combination of the feature extraction process together with the verification process. Therefore, in this research the DWT will be used with various types of mother wavelet in various order to get a comparative analysis. From this comparative analysis, the best mother wavelet and the order of the mother wavelet can be determine for use in an ECG biometric verification system.

In ECG biometric verification, the feature vectors need to be classified into multiple classes as an individual's ECG signal will be compared to multiple individuals in a collected database. A linear classifier is only able to distinguish data into two classes while non-linear classifier can classify data into multiple classes which are more rational to use in ECG biometric applications. Considering the classifier discussed in this literature, the linear classifier in [2] yields poor performance when compared to the non-linear classifier in [6][7][18] which yield better performances. With the use of a non-linear classifier, the biometric verification can be developed together with the feature extraction. The ECG biometrics studies are summarized in Table 2.1.

Table 2.1: Summary of studies in ECG biometrics.

Papers	Database	Feature Extraction Method	Verification Method	Performance	Comments
Hegde *et. al* (2011) [12]	PhysioNet QT, MIT-BIH	Radon transform	Correlation coefficient	FAR of 2.19% and FRR of 0.128%	Proposed an image based feature vector.
Safie *et. al* (2011) [15]	PTB	PAR	Euclidean distance	EER of 0.1915	Studied on both healthy and unhealthy subject.
Safie *et. al* (2011) [14]	PTB	PAW	Euclidean distance	EER of 0.1801	Proposed a unique way of extracting feature vector.
Agrafioti (2011) [2]	Self-collected	Autocorrelation	LDA	GAR of 92.3%	Proposed ways to capture emotion from ECG signal to use as ECG biometric verification.
Belgacem *et. al* (2012) [6]	MIT-BIH, ST-T, NSR, PTB and self-collected	DWT	Random Forest	GAR of 100%	Uses variety of database and emphasized on the use of DWT as feature extractor.

Table 2.1: Summary of studies in ECG biometrics (continued).

Papers	Database	Feature Extraction Method	Verification Method	Performance	Comment
Hamdi ***et. al*** **(2014)** [13]	Self-collected	Slopes and angles in ECG signal	MLP Neural network	GAR of 96.44%.	Complex classifier used and proposed slope and angle parameter as feature vector.
Bashar ***et. al*** **(2015)** [17]	PTB	MLSP	Euclidean distance	GAR of 91.67%	Proposed an image based feature vector.
Akhter ***et. al*** **(2016)** [16]	Self-collected	Statistical, spectral, wavelet and nonlinear techniques	K-Nearest Neighbor	GAR of 82.22%	Proposed HRV for used in biometric verification system.
Hejazi ***et. al*** **(2016)** [18]	Self-collected	Autocorrelation & KPCA	Gaussian OAA SVM	FRR of 4.83% and FAR of 3.5%	Emphasized on the kernel trick used for feature extraction as well as for classifier.
Ramli ***et. al*** **(2016)** [7]	Self-collected	DWT	SVM	GAR of 97%, FAR of 1% and EER of 2%	Proposed a portable ECG detection kit for used in ECG biometric verification.

2.6 Summary

This chapter presented a discussion on the research of ECG biometric verification system developed based on combination of different feature extraction method and classifiers used for verification. From the researches in biometric verification, the uses of non-fiducial based feature extraction have more advantages as compared to the fiducial based feature extraction. The DWT is one of the non-fiducial based feature extraction method that shows better performance but there is still ambiguity in the performance when different mother wavelets are used. Furthermore, the literature also shows that the non-linear classifier is giving better performance result as compared to the linear classifier. This shows that the SVM is one of the best non-linear classifier to be used in the verification system. Therefore, an ECG biometric verification system will be developed using the DWT as feature extraction and the SVM classifier for verification.

Besides that, further studies are required in the DWT as a feature extractor due to the different level of decomposition and mother wavelet may play an important role in extracting the features accurately. The different types and mother wavelet will also impact the performance of the ECG biometric verification system. Hence, a comparative analysis should be conducted to show the performance of the biometric verification system. This comparative study should use different types and orders of mother wavelets to determine which mother wavelets will provide the best performance for an ECG biometric verification system.